

# *An Examination of New Paradigms for Spline Approximations*

Volume 111

Number 2

March-April 2006

**Christoph Witzgall, David E. Gilsinn, and Marjorie A. McClain**

National Institute of Standards and Technology,  
Gaithersburg, MD 20899-8910

christoph.witzgall@nist.gov  
david.gilsinn@nist.gov  
marjorie.mcclain@nist.gov

Lavery splines are examined in the univariate and bivariate cases. In both instances relaxation based algorithms for approximate calculation of Lavery splines are proposed. Following previous work Gilsinn, et al. [7] addressing the bivariate case, a rotationally invariant functional is assumed. The version of bivariate splines proposed in this paper also aims at irregularly spaced data and uses Hsieh-Clough-Tocher elements based on the triangulated irregular network (TIN) concept. In this paper, the univariate case, however, is investigated in greater detail so as to further the understanding of the bivariate case.

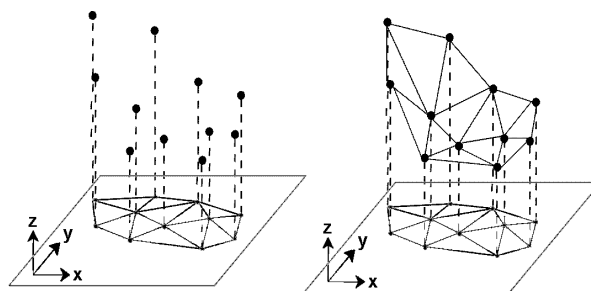
**Key words:** bivariate splines; curve fitting; Delaunay triangulation; Gauss-Seidel iteration; Hsieh-Clough-Tocher elements; irregular data; Lavery splines; non-oscillatory splines; point clouds; surface fitting; thin beam; thin plate; triangulated irregular networks.

**Accepted:** December 14, 2005

**Available online:** <http://www.nist.gov/jres>

## 1. Introduction

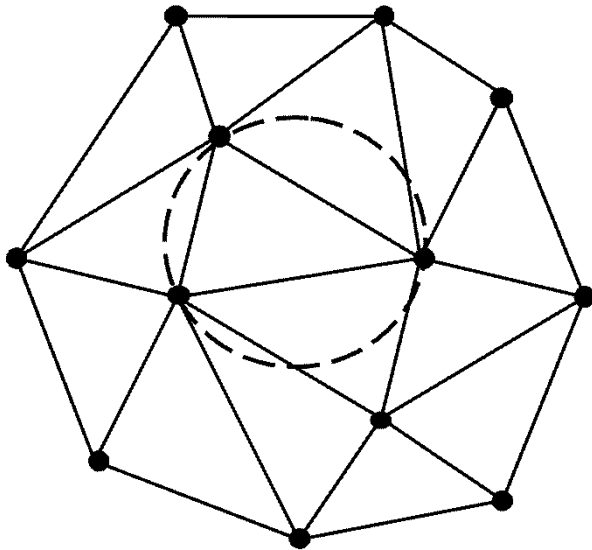
Over the last decade the task of representing very large “point clouds” by meshed surfaces has arisen in many applications. Such point clouds may have been generated from terrain surveys or from LADAR (LAsER Direction And Ranging) scans of objects. Work has centered around the concept and application of Triangulated Irregular Network (TIN). Here the “foot prints”  $(x_i, y_i)$  of data points  $(x_i, y_i, z_i)$  are “triangulated”, that is, the area of interest is tiled by non-overlapping triangles defining a piecewise triangular “TIN-surface” in space as each triangle is lifted according to the elevations  $z_i$  at its corners as indicated in Fig. 1. Most often, the triangulation used is the “Delaunay” triangulation, characterized by the “empty circle property” in which the circumcircle of each triangle does not contain any triangle corners in its interior (see Fig. 2). Figure 3 features such a Delaunay triangulation, under-



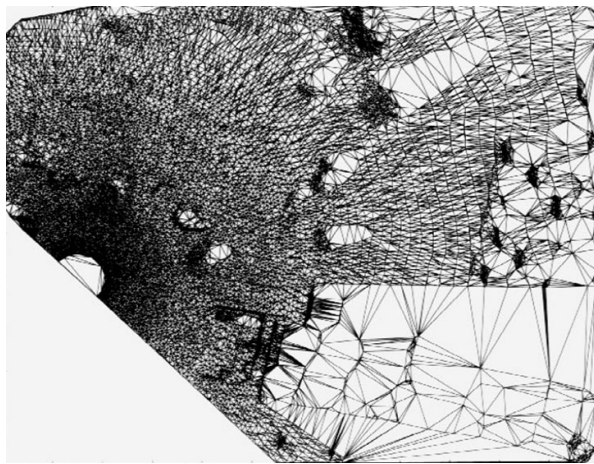
**Fig. 1.** TIN meshing: Triangulated surface over a triangulation in the footprint plane.

lying a TIN, created from a LADAR scan of a construction site on NIST grounds. The construction of the surface from adjacent patches or “elements” is basically a “Finite Element” technique.

TIN constructs utilize the actual data points directly even if they are distributed irregularly rather than refer-



**Fig. 2.** Delaunay triangulation: Circumcircles of triangles do not contain vertices in their interior.



**Fig. 3.** A TIN created from the LADAR scan of terrain.

ring to approximations by regular rectangular grids. They adapt naturally to disparate data densities. Importantly, they support strategic selection of data points to be used as the support points for a TIN surface.

Conceptually, TINs define which points may be considered “neighbors” of each other. This proximity concept based on a direct neighbor criterion differs from concepts based on a fixed distance cutoff in that it adjusts automatically to differing data densities such as in data collected by ground based LADAR. Figure 3 exemplifies the dramatic differences in densities encountered in ground-based LADAR scans. Also, because of the neighbor relation provided by a TIN, the

utility provided by that construct is not restricted to surface generation, but extends to data editing and analysis.

The TIN surface, as defined according to Fig. 1, is  $C^0$ , that is, it is continuous but usually not smooth: it is typically not differentiable along triangle edges, since the plane of any of the spatial triangles constituting the surface generally differs from the planes of its adjacent triangles. The TIN concept of triangulation-based surface generation is frequently understood to imply the use of such planar “elements” resulting in a non-smooth, piecewise linear TIN surface. It should be emphasized, however, that the TIN concept also supports the use of non planar, that is, curved elements for a smooth or  $C^1$  surface. Planar elements offer some advantages besides simplicity and a certain robustness to be discussed below. Visualization still requires piecewise linear representations in order to delineate hidden surfaces. Volume computations similarly tend to be either grid-based or based on piecewise linear surfaces for ease of computation. In those cases, a smooth surface would have to be discretized or approximated by a piecewise linear surface. This then begs the question, why not use planar elements in the first place?

Without any doubt, however, there are many instances, in which a smooth surface representation would provide definite advantages:

- More accuracy could be achieved without increasing memory requirements because curved elements encode more geometric information than planar elements.
- Clusters of coplanar points are avoided when sampling the surface at a discrete set of foot prints such as regular grid points.
- When simulating movement over a terrain surface, a smooth ride is preferable. A ride over a piecewise linear surface is necessarily bumpy at transitions from one triangle to another.
- In many applications, several point clouds which were collected with reference to different coordinate systems need to be “registered”, that is, combined within a common coordinate system. Some common registration methods, such as point-to-surface iterative closest point (ICP) proposed by Besl [2], require minimizing the deviations of points in one point cloud from the surface representation of another point cloud. This minimization process should work much better if that surface were smooth.

In computer aided design (CAD), design-defined objects are routinely represented by surfaces that are  $C^1$  or even  $C^2$ . Terrain representation, however, poses a

different challenge: The location of “crease lines” or “break lines”, along which the actual terrain surfaces are not differentiable, are usually not known ahead of time, whereas in a CAD environment break lines tend to be specified as part of the design. Unspecified break lines, on the other hand, along with actual verticalities, tend to give rise to spurious oscillation and, what may be called, “Gibbs phenomena” in analogy to the phenomenon well known from the theory of Fourier series. TIN surfaces based on planar elements, on the other hand, are more robust, and can be constructed so as to automatically represent and, if necessary, report break lines. Susceptibility to spurious oscillations and Gibbs phenomena are one of the reasons why the terrain modeling community has been slow to accept smooth surfaces. There has thus been a long quest for “non-oscillatory splines” which would obviate this pesky conundrum.

In 1994, Lavery [11,12,13] (see also Gilsinn and Lavery [8,9,10]) proposed successful paradigms for univariate as well as bivariate non-oscillatory splines, which could be used for representing 2D or 3D data sets, respectively. Lavery introduced the term “ $L_1$  splines” for his brand of nonoscillatory splines. The term  $L_1$  splines is, however, frequently misinterpreted as minimizing an  $L_1$  measure-of-fit when approximating a point set by, say, classical splines. For this reason, we prefer the term “Lavery splines”.

Classical splines are characterized by their minimizing energy functionals. Lavery splines, on the other hand, minimize different functionals. In the bivariate case, especially, the computational effort of minimizing these functionals, however, exceeds the effort required by the classical approach by an order of magnitude. For the bivariate case we have, therefore, considered an approximation to the calculation of Lavery splines, along with a prior modification of the functional proposed by Lavery for the bivariate case. This modified bivariate functional is still an extension of Lavery’s functional for univariate functions, but it extends a different aspect of the latter. It also is invariant under planar rotations of the coordinate system, which Lavery’s bivariate functional is not. We will present preliminary results for both univariate and bivariate nonoscillatory splines.

The main thrust of this paper, however, remains to illustrate the utility and performance of bivariate non-oscillatory splines, building on the previous study by Gilsinn et al. [7], and also on an early terrain modeling study by Mandel et al. [18]. Both studies employ TIN techniques in conjunction with the “reduced Hsieh-Clough-Tocher (rHCT)” element, the approach was

pioneered by Lawson [14,15] and is followed in this work. It involves specifying elevations  $z_i$  and two slopes  $z_{ix}, z_{iy}$  at each triangle corner or vertex  $v_i = (x_i, y_i)$  in a triangulation, and filling in, at each triangle, the thereby defined rHCT elements, results in a smooth surface over an entire TIN.

The classical spline approach to, say, interpolating the elevations at vertices of a triangulation would be to prescribe the elevations  $z_i$ , and select the remaining parameters, namely the partial slopes  $z_{ix}, z_{iy}$ , so as to minimize some energy functional. This was indeed the approach taken in the work by Mandel et al. [18] The task to be accomplished there was to represent terrain given by digitized contour lines. The elevations at the data points were thus determined by the elevation of the contour line on which the data point was located. In the end, elevations were to be evaluated at the points of a square  $900 \times 900$  grid. At the time, — 1985 — visualization procedures that are now routine were not yet commonly available. A display of the results had to wait another year. So assessing the quality of the representation was restricted to manual spot checking. It thus took until a day before a scheduled presentation that, to our dismay, huge oscillations were detected. Fortunately, it turned out that it was not due to a problem with the method but was caused by an error in the data that had been provided: the elevation of a single contour line had been tagged 100 feet too high, causing the disruption. In this instance, the oscillations proved to be actually beneficial in that they uncovered data errors.

In Sec. 2 we discuss aspects of univariate Lavery splines in order to shed light on related issues for bivariate splines, which are addressed in Sec. 3. The algorithm proposed here for bivariate non-oscillatory splines requires solving large sparse systems of linear equations. Experience was gathered concerning the performance of the well known Gauss-Seidel algorithm.

## 2. Univariate Spline Interpolation

While the emphasis of this work is on bivariate splines for surface generation, an examination of univariate splines is taken up in order to highlight some of the issues pertaining to splines in general. In the univariate case, cubic “spline functions” are most commonly used and are considered here. They form a linear space

$F$

of piecewise cubic  $C^1$  functions  $f(x)$  defined locally over intervals between “knots”

$$x_0 < x_1 < \dots < x_n,$$

that is, they consist of cubic polynomials

$$f_i(x), x \in [x_{i-1}, x_i], i = 1, \dots, n.$$

Adjacent cubic polynomials are required to assume the same values  $y_i$  at common interior knots,

$$y_i = f_i(x_i) = f_{i+1}(x_i).$$

This ensures continuity of the complete spline function  $f(x)$  over the entire interval

$$I = [x_0, x_n].$$

In addition, the polynomials are to assume the same slopes

$$m_i = f'_i(x_i) = f'_{i+1}(x_i).$$

The spline functions  $f(x)$  are thus continuously differentiable, that is, they belong to class  $C^1$ . In what follows, the linear spaces

$$F', F''$$

of first and second derivatives of spline functions are also considered, in spite of the fact that, at common knots, the second derivatives of adjacent cubic polynomials may not agree, so that the spline functions  $f(x) \in F$  are generally not twice differentiable at such knots. However, they are twice differentiable everywhere but on this set of measure zero. For the purposes of integration below, it does not matter that the function  $f''(x)$  may not be defined for those arguments.

Each of the constituent cubic polynomials  $f_i(x)$  is uniquely determined by the values  $y_{i-1}, y_i$  at the knots  $x_{i-1}, x_i$  and the slopes  $m_{i-1}, m_i$  at those locations (Fig. 4), in fact, the polynomial is linear in these parameters. The entire function  $f(x)$  is thus uniquely determined by its values and slopes at the knots, and it, too, depends linearly on these parameters, so that the space  $F$  is isomorphic to the  $2(n+1)$ -dimensional vector space of values  $y_i$  and slopes  $m_i, i = 0, \dots, n$ .

We now turn to the task of interpolation. Here the values  $y_i$  at the knots  $x_i$  are fixed and specified. Given a particular specification of values  $y_i$ , a corresponding

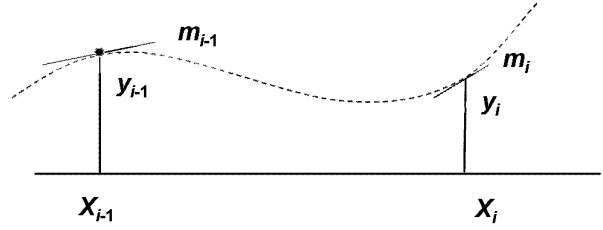


Fig. 4. Hermite cubic polynomial determined by coordinates and slopes at end points.

“interpolating spline function” depends only on the parameters  $m_i$ :

$$f(x) = f(m_0, m_1, \dots, m_n; x).$$

Collectively, these functions define affine manifolds or flats,

$$S, S', S'',$$

within the linear spaces  $F, F', F''$ , respectively. That is, if

$$m_i = \frac{m_i^{(1)} + m_i^{(2)}}{2}, i = 0, 1, \dots, n,$$

then it follows that correspondingly

$$\begin{aligned} f(x) &= f(m_0, m_1, \dots, m_n; x) \\ &= \frac{f(m_0^{(1)}, m_1^{(1)}, \dots, m_n^{(1)}; x) + f(m_0^{(2)}, m_1^{(2)}, \dots, m_n^{(2)}; x)}{2}. \end{aligned}$$

The question then becomes, how to select slopes  $m_i$  so as to achieve a “satisfactory” interpolation. That selection is generally made by minimizing a functional on the affine space  $S$  defined by an integral. In this work, we reserve the term “spline” — as opposed to “spline function”, for the results of such a minimization.

## 2.1 Paradigms for Univariate Splines

“Classical splines” are uniquely defined as those interpolating cubic spline functions which are (i)  $C^2$ , that is, twice differentiable, such that

$$(1) \quad f''_i(x_i) = f''_{i+1}(x_i), i = 1, \dots, n-1$$

holds at interior knots, and for which (ii) the second derivative vanishes

$$(2) \quad f''(x_0) = f''(x_n) = 0$$

at the two exterior knots. Holladay proved early on (see Ahlberg [1]) that classical splines are also defined as the unique minimizers of the “thin beam energy”

$$(3) \quad E(f) = \int_{x_0}^{x_n} f''(x)^2 dx$$

over the affine space  $S$  of all  $C^1$  interpolating spline functions.

Condition (2) is familiar to structural engineers as the vanishing of the second derivative at the “free end” of a beam. It is, however, remarkable that this energy minimization enforces also a higher level of compatibility across knots (1) so that the minimizing  $C^1$  spline functions are, in fact, in class  $C^2$ . On the other hand, this stiffness contributes to the tendency of classical splines to produce spurious oscillations, Gibbs phenomena and undesirable inflections.

Several attempts have been made to avoid these problems. Taking a clue from mechanics, Schweikert [20] and Cline [3] have introduced “tension splines”, where the arclength of the spline function is made part of the defining minimization. Reinsch (see Stoer and Bulirsch [21]) moved to “exponential splines” by adding a multiple of the square of first derivatives to the integrand in (3). Those efforts were only partially successful. Drawbacks include the need to specify an additional parameter in order to balance conflicting minimization requirements, and the fact that these techniques are not readily generalized to the bivariate case.

Lavery splines, on the other hand, appear to avoid such shortcomings. In Figs. 5 and 6 we compare Lavery splines against classic splines. We are particularly impressed with the performance of Lavery splines for the example in Fig. 8, as compared to the example in Fig. 7. What is the secret of Lavery splines? How are they defined as opposed to classical splines?

Lavery defines his nonoscillatory splines, in essence, as minimizing the integral of the absolute value rather than the square of the second derivative of a spline function:

$$(4) \quad e(f) = \int_{x_0}^{x_n} |f''(x)| dx$$

over the given affine space  $S$  of interpolating spline functions.

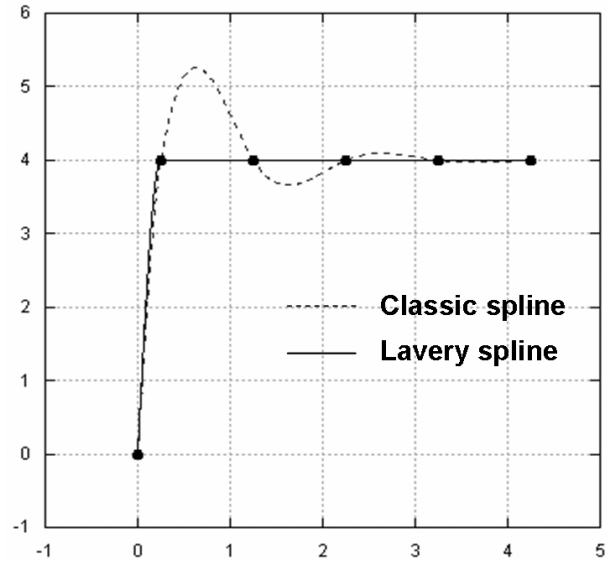


Fig. 5. Lavery spline vs. classic spline: Example 1.

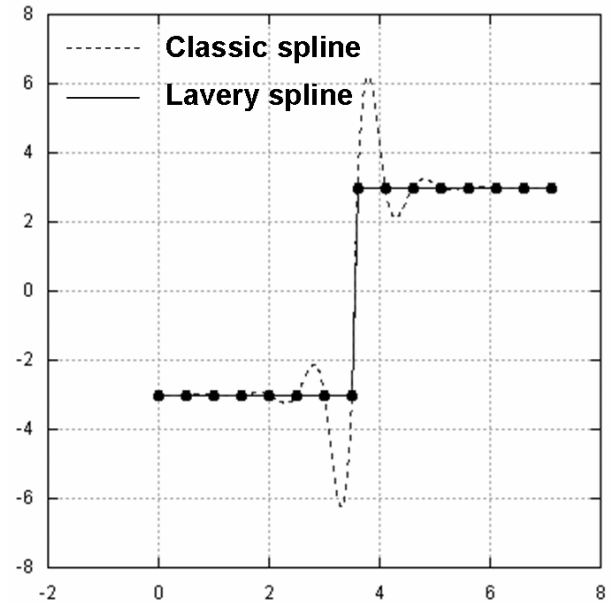


Fig. 6. Lavery spline vs. classic spline: Example 2.

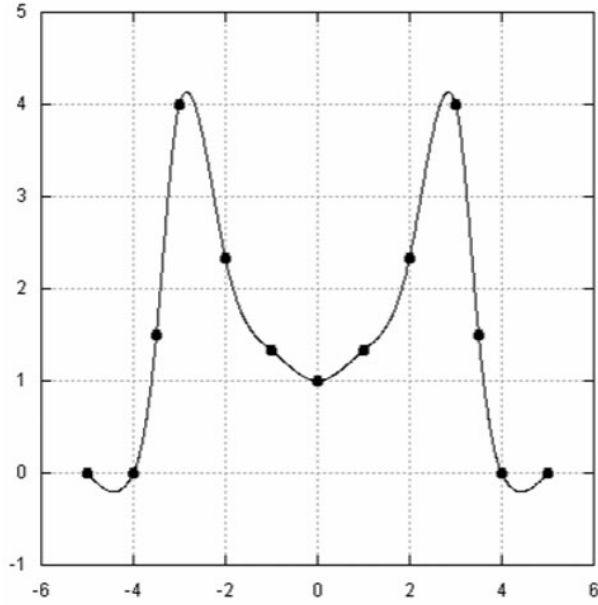


Fig. 7. Classic splines produce unnecessary variations in curvature.

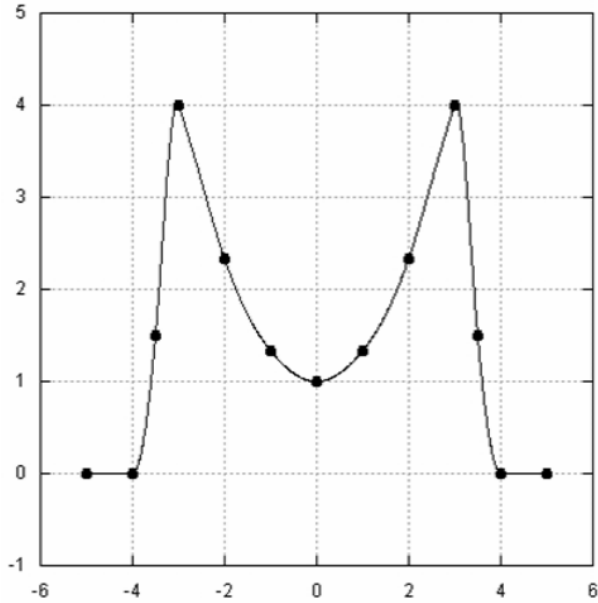


Fig. 8. Lavery spline has smooth curvature transitions.

## 2.2 Expressing Holladay and Lavery Integrals

For the purposes of this paper, we will refer to the integrals (3) and (4) as the “Holladay integral” and the “Lavery integral”, respectively. The goal of this section is to derive expressions for the values of these integrals in terms of the slopes  $m_i$ ,  $i = 1, \dots, n$ , at the knots of the spline function  $f(x)$  under consideration. Both are the sums of the corresponding integrals of the second derivatives of the individual cubic polynomials  $f_i(x)$ , which constitute the spline function  $f(x)$ .

As pointed out in Sec. 2.1, each such cubic polynomial is uniquely defined by its end points  $(x_{i-1}, y_{i-1})$ ,  $(x_i, y_i)$  and its end slopes  $m_{i-1}$ ,  $m_i$ . The various formulas describing this polynomial are commonly referred to as “Hermite” formulas. For the versions used here, we introduce the quantities

$$\Delta_i = x_i - x_{i-1}$$

and

$$M_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}} = \frac{y_i - y_{i-1}}{\Delta_i},$$

where  $M_i$  represents the slope of the straight line between end points. Instead of referring to the variable  $x$  directly, the following formulas are in terms of the weights

$$(5) \quad \begin{aligned} \lambda_i &= \lambda_i(x) = \frac{x_i - x}{x_i - x_{i-1}} = \frac{x_i - x}{\Delta_i}, \\ \mu_i &= \mu_i(x) = \frac{x - x_{i-1}}{x_i - x_{i-1}} = \frac{x - x_{i-1}}{\Delta_i}, \end{aligned}$$

where  $\lambda_i + \mu_i = 1$ , and  $\lambda_i, \mu_i \geq 0$ , for  $x$  in the interval  $[x_{i-1}, x_i]$ . Such weights are often referred to as “barycentric coordinates” or, in the bivariate case, as “triangle coordinates”. With these conventions, we find, for instance,

$$(6) \quad f_i(x) = \lambda_i y_{i-1} + \mu_i y_i + \lambda_i^2 \mu_i (m_{i-1} - M_i) \Delta_i - \lambda_i \mu_i^2 (m_i - M_i) \Delta_i.$$

Furthermore, by definition (5),

$$dx = -\Delta_i d\lambda_i = +\Delta_i d\mu_i,$$

the chain rule yields, using  $(\lambda_i + \mu_i)^2 = 1$ ,

$$f'_i(x) = 6\lambda_i\mu_i M_i + (\lambda_i^2 - 2\lambda_i\mu_i)m_{i-1} + (\mu_i^2 - 2\lambda_i\mu_i)m_i.$$

An alternate expression for the first derivative is readily derived:

$$(7) \quad f'_i(x) = \lambda_i m_{i-1} + \mu_i m_i + \lambda_i \mu_i D_i.$$

Here the quantity

$$(8) \quad D_i = 6M_i - 3m_{i-1} - 3m_i = 6 \left( M_i - \frac{m_{i-1} + m_i}{2} \right),$$

vanishes if and only if the polynomial  $f(x)$  has degree less than three. Because by (5)

$$\lambda_i \mu_i = \left( \frac{x_i - x}{\Delta_i} \right) \left( \frac{x - x_{i-1}}{\Delta_i} \right)$$

$-D_i$  is seen as the lead coefficient of  $f'_i(x)$  as expressed in  $x$ . Consequently  $-1/3D_i$  is the lead coefficient of  $f''_i(x)$ .  $D_i < 0$  indicates that the function is concave up to its inflection point and convex thereafter. Conversely,  $D_i > 0$  indicates that convexity precedes concavity in the direction of the  $x$ -axis. The quantity  $D_i$  will play a major role in what follows. The same is true for the quantities  $U_i$ ,  $V_i$  in the expression

$$(9) \quad f''_i(x) = \frac{2}{\Delta_i} (\lambda_i U_i - \mu_i V_i),$$

where

$$(10) \quad \begin{aligned} U_i &= 3M_i - 2m_{i-1} - m_i = +\frac{\Delta_i}{2} f''_i(x_{i-1}) \\ V_i &= 3M_i - m_{i-1} - 2m_i = -\frac{\Delta_i}{2} f''_i(x_i). \end{aligned}$$

The coefficients  $U_i$ ,  $V_i$  thus relate to the two-sided second derivatives of  $f_i(x)$  at knots  $x_i$ . Note that

$$D_i = U_i + V_i$$

and also that, in view of (5),

$$\int_{x_{i-1}}^{x_i} \lambda_i^2 dx = \frac{\Delta_i}{3}, \quad \int_{x_{i-1}}^{x_i} \lambda_i \mu_i dx = \frac{\Delta_i}{6}, \quad \int_{x_{i-1}}^{x_i} \mu_i^2 dx = \frac{\Delta_i}{3}.$$

An expression for the thin beam energy of the individual polynomial  $f_i(x)$  is now readily derived:

$$(11) \quad E(f_i) = \frac{4}{\Delta_i} [m_{i-1}^2 + m_{i-1}m_i + m_i^2 - 3M_i(m_{i-1} + m_i) + 3M_i^2].$$

The full Holladay integral is the inhomogeneous quadratic function of the variables  $m_i$ , arrived at by adding the energies  $E(f_i)$  of all partial functions  $f_i(x)$ . Introducing the vector of slopes

$$\tilde{\mathbf{m}} = (m_0, m_1, \dots, m_n)^T,$$

we have in matrix notation

$$E(f) = \tilde{\mathbf{m}}^T \mathbf{H} \tilde{\mathbf{m}} - 12\tilde{\mathbf{M}}^T \tilde{\mathbf{m}} + 12c,$$

where

$$\mathbf{H} = \begin{pmatrix} \frac{4}{\Delta_1} & \frac{2}{\Delta_1} & & & & \\ \frac{2}{\Delta_1} & \frac{2}{\Delta_1} + \frac{2}{\Delta_2} & \frac{2}{\Delta_2} & & & \\ & \frac{2}{\Delta_2} & \frac{2}{\Delta_2} + \frac{2}{\Delta_3} & \frac{2}{\Delta_3} & & \\ & & \vdots & \vdots & \vdots & \\ & & & \frac{2}{\Delta_{n-1}} & \frac{2}{\Delta_{n-1}} + \frac{2}{\Delta_n} & \frac{2}{\Delta_n} \\ & & & & \frac{2}{\Delta_n} & \frac{4}{\Delta_n} \end{pmatrix},$$

$$\tilde{\mathbf{M}} = \begin{pmatrix} \frac{M_1}{\Delta_1} \\ \frac{M_2}{\Delta_2} + \frac{M_1}{\Delta_1} \\ \frac{M_3}{\Delta_3} + \frac{M_2}{\Delta_2} \\ \vdots \\ \frac{M_n}{\Delta_n} + \frac{M_{n-1}}{\Delta_{n-1}} \\ \frac{M_n}{\Delta_n} \end{pmatrix}$$

and

$$c = \left(\frac{M_1}{\Delta_1}\right)^2 + \left(\frac{M_2}{\Delta_2}\right)^2 + \dots + \left(\frac{M_n}{\Delta_n}\right)^2.$$

Minimizing this expression for the Holladay energy integral is to solve the linear system of equations

$$(12) \quad 2\mathbf{H}\tilde{\mathbf{m}} = \mathbf{M}$$

or

$$\begin{aligned}
\frac{2}{\Delta_1} m_0 + \frac{1}{\Delta_1} m_1 &= \frac{3}{\Delta_1} M_1 \\
\frac{1}{\Delta_1} m_0 + \left( \frac{2}{\Delta_1} + \frac{2}{\Delta_2} \right) m_1 + \frac{1}{\Delta_2} m_2 &= \frac{3}{\Delta_1} M_1 + \frac{3}{\Delta_2} M_2 \\
\frac{1}{\Delta_2} m_1 + \left( \frac{2}{\Delta_2} + \frac{2}{\Delta_3} \right) m_2 + \frac{1}{\Delta_3} m_3 &= \frac{3}{\Delta_2} M_2 + \frac{3}{\Delta_3} M_3 \\
&\dots = \dots \\
\frac{1}{\Delta_n} m_{n-1} + \frac{2}{\Delta_n} m_n &= \frac{3}{\Delta_n} M_n
\end{aligned}$$

The first equation, in fact, may be restated by (10) as

$$\frac{1}{\Delta_1} U_1 = \frac{1}{2} f_1''(x_0) = 0,$$

that is, as the requirement (2) that second derivatives vanish at the first knot. The last equation reflects the corresponding requirement at the last knot. The second equation is equivalent to

$$\frac{1}{\Delta_1} V_1 + \frac{1}{\Delta_2} U_2 = \frac{1}{2} f_1''(x_1) - \frac{1}{2} f_2''(x_1) = 0,$$

implying second order differentiability at the interior knot  $x_1$ . The remaining equations similarly enforce compatibility of the one-sided second derivatives at the remaining interior knots. This confirms the result of Holladay.

The reader should note that the system (12) for classical splines differs from the linear system mostly offered in the current literature. There the second order differentiability of the classical splines is already assumed and the spline functions are formulated in terms of those second order derivatives,  $n_i = f''(x_i)$ ,  $i = 0, \dots, n$ . However for weighted classical splines, to be encountered in Sec. 2.4.1, second order differentiability no longer holds, and a weighted version of the linear system (12) needs to be considered.

### 2.2.1 Expressing the Lavery Integral

As to the integral of the absolute value of the second derivative, it is readily available in the case that the derivative does not change sign inside the subinterval:

$$\begin{aligned}
e(f_i) &= \int_{x_{i-1}}^{x_i} |f_i''(x)| dx = \left| \int_{x_{i-1}}^{x_i} f_i''(x) dx \right| \\
&= |f_i'(x_i) - f_i'(x_{i-1})| = |m_i - m_{i-1}|.
\end{aligned}$$

This includes the case in which the polynomial is of lesser degree than cubic, and thus has constant second

derivatives. This case is signaled by the vanishing of the quantity  $D_i$  introduced earlier in (8).

The function  $f''(x)$ , however, is a linear function in  $x$  and, unless constant, changes sign at some location  $\hat{x}$ , which also marks the location of the inflection point of  $f(x)$ . Suppose this location falls into the interior of the subinterval:

$$x_{i-1} < \hat{x} < x_i.$$

Then the integral has to be calculated as the sum of two integrals of linear functions:

$$e(f_i) = \left| \int_{x_{i-1}}^{\hat{x}} f_i''(x) dx \right| + \left| \int_{\hat{x}}^{x_i} f_i''(x) dx \right|.$$

The integrals between the absolute value bars are of opposite signs, so that the total integral can be written as the absolute value of a difference of integrals

$$e(f_i) = \left| \int_{x_{i-1}}^{\hat{x}} f_i''(x) dx - \int_{\hat{x}}^{x_i} f_i''(x) dx \right|.$$

These integrals can be separately evaluated as differences of slopes. Let

$$\hat{m}_i = f'(\hat{x})$$

denote the inflection slope, then — in the case of an interior inflection —

$$(13) \quad e(f_i) = |2\hat{m}_i - m_{i-1} - m_i|.$$

The quantities  $U_i$ ,  $V_i$  depend linearly on the two slopes  $m_{i-1}$ ,  $m_i$ . Conversely, those slopes can be expressed in terms of  $U_i$ ,  $V_i$ .

$$(14) \quad m_{i-1} = M_i - \frac{2}{3} U_i + \frac{1}{3} V_i, \quad m_i = M_i + \frac{1}{3} U_{i-1} - \frac{2}{3} V_i.$$

The next step is to express the inflection slope in terms of  $U_i$ ,  $V_i$ . To this end, we express the inflection argument by its barycentric weights:

$$(15) \quad \hat{x} = \hat{\lambda}_i x_{i-1} + \hat{\mu}_i x_i,$$

From the expression (10) for  $f_i''(x)$  it follows that

$$\hat{\lambda}_i = \frac{V_i}{D_i}, \quad \hat{\mu}_i = \frac{U_i}{D_i},$$

where the denominator  $D_i = U_i + V_i = 6M_i - 3m_{i-1} - 3m_i$  has already been encountered in (8) as a quantity that vanishes if and only if the polynomial in question is



parabolic or linear. Thus  $D_i = 0$  leads back to the previous case of no sign changes by the second derivative.

The weights  $\hat{\lambda}_i, \hat{\mu}_i$  may now be inserted into the expression (7) for the derivative  $f'_i(x)$ . This gives

$$\hat{m}_i = \frac{V_i}{D_i} m_{i-1} + \frac{U_i}{D_i} m_i + \frac{V_i U_i}{D_i}.$$

Substituting for  $m_{i-1}, m_i$  according to (14) yields

$$\hat{m}_i = M_i + \frac{U_i^2 + V_i^2 - V_i U_i}{3D_i}$$

as well as

$$(16) \quad e(f_i) = \frac{U_i^2 + V_i^2}{3|D_i|} = \frac{\hat{\mu}_i^2 + \hat{\lambda}_i^2}{3|D_i|}.$$

in view of (19). In terms of the slopes  $m_{i-1}, m_i$ :

$$e(f_i) = \left| \frac{18M_i^2 - 18M_i(m_{i-1} + m_i) + 5m_{i-1}^2 + 8m_{i-1}m_i + 5m_i^2}{6M_i - 3m_{i-1} - 3m_i} \right|.$$

The above expressions for  $e(f_i)$  are also valid if the inflections occur at the ends  $x_{i-1}, x_i$  of the interval of definition, reducing to  $e(f_i) = |m_i - m_{i-1}|$ , in accordance with earlier results.

We are now ready to examine the full Lavery integral. At first blush, all that remains to be done is to sum over the partial integrals  $e(f_i)$  in their various forms. We will show, however, that many terms of the expressions (13) cancel each other out as these partial integrals are added together. To this end, we distinguish five separate kinds of polynomials  $f(x)$  depending on their behavior in the interior of the interval between its knots,  $x_{i-1} < x < x_i$ :

- “Linear”; here  $f''_i(x) = 0$  throughout and  $e(f_i) = 0$
- “Convex”; here  $f''_i(x) > 0$  in the interior of the interval  $[x_{i-1}, x_i]$  and  $e(f_i) = m_i - m_{i-1}$
- “Concave”; here  $f''_i(x) < 0$  in the interior of the interval  $[x_{i-1}, x_i]$  and  $e(f_i) = m_{i-1} - m_i$
- “Convex-concave”; here  $f''_i(x) > 0$  for  $x < \hat{x}$ ,  $f''_i(x) < 0$  for  $x > \hat{x}$  and  $e(f_i) = +\hat{m} - m_{i-1} - m_i$ , also  $D_i > 0$
- “Concave-convex”; here  $f''_i(x) < 0$  for  $x < \hat{x}$ ,  $f''_i(x) > 0$  for  $x > \hat{x}$  and  $e(f_i) = -\hat{m} + m_{i-1} + m_i$ , also  $D_i < 0$

The last two categories are the ones with an interior inflection point  $\hat{x}$  and inflection slope  $\hat{m}$ .

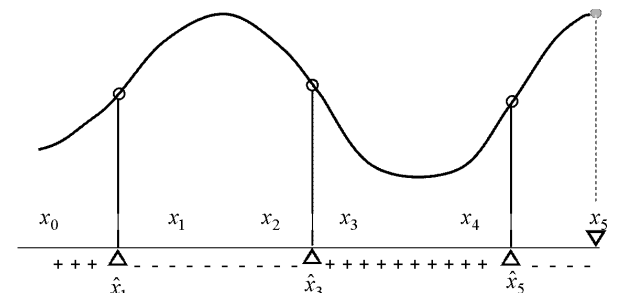
The interior inflection points considered so far may not be the only inflection points of the spline function  $f(x)$ . Inflections occur also at knots  $x_i$ ,  $0 < i < n$  if a concave or convex-concave polynomial is followed by a convex or convex-concave polynomial and, analogously, if a convex or concave-convex polynomial is followed by a concave or concave-convex polynomial. We refer to such knots as “inflection knots”. To make matters more complicated, however, inflection may occur along an entire stretch of consecutive linear polynomials of equal slope, the inflection slope in this case, provided there are adjacent nonlinear polynomials at both ends of such a stretch exhibiting the same convexity/concavity pattern that would cause an inflection at a knot. In this case, we choose an arbitrary knot, say, the leftmost one in the linear stretch, as the inflection knot representing the inflection.

Clearly, slopes at interior knots that are not inflections cancel out as the expressions (13) for the Lavery integrals  $e(f_i)$  for the partial spline functions  $f_i(x)$  are added up. See Fig. 9 for an example. This leads to

**Observation A:** *The Lavery integral of a cubic spline function is the absolute value of an alternating sum of the inflection slopes and of the end slopes. Let*

$$\hat{m}_1, \hat{m}_2, \hat{m}_3, \dots, \hat{m}_L,$$

*be the sequence of all inflection points identified above, sorted from left to right by their location. Then*



$$\int_{x_0}^{x_5} |f''(x)| dx = \left| \int_{x_0}^{\hat{x}_1} - \int_{\hat{x}_1}^{\hat{x}_3} + \int_{\hat{x}_3}^{\hat{x}_5} - \int_{\hat{x}_5}^{x_5} \right|$$

$$= |-m_0 - 2\hat{m}_1 - 2\hat{m}_3 + 2\hat{m}_5 - m_5|$$

**Fig. 9.** Lavery integral as expressed by inflection slopes and end slopes.

$$e(f) = \left| m_0 + 2 \sum_{l=1}^L (-1)^l \hat{m}_l - (-1)^L m_n \right|.$$

(Note that the indices  $l$  of  $\hat{m}_l$  do not refer to the interval in which they are located).

### 2.3 Properties of Lavery Splines

In this section, we gather some information about Lavery Splines, namely, interpolating spline functions in the affine space  $S$  which minimize their respective Lavery integrals. A first general observation concerns the convexity of the Lavery integral.

The Holladay and Lavery integrals (3) and (4) of a piecewise cubic spline function,

$$f(x) = f(m_0, m_1, \dots, m_n; x)$$

are functions of the slope specifications  $m_i$ :

$$E(f) = E(m_1, m_2, \dots, m_n), \quad e(f) = e(m_1, m_2, \dots, m_n).$$

The quadratic function  $E(f)$  representing the Holladay integral can be shown to be positive definite and, therefore, strictly convex. Its minimum is unique on the affine manifold  $S$  interpolating spline functions. The restriction to  $S$  is, of course, necessary as the value of  $E(f)$  would not change if the spline function  $f(x)$  were modified by adding a linear function in  $x$ . Adding a non-zero linear function to the function  $f(x)$  would not preserve its interpolation property.

#### 2.3.1 Convexity

Next, we establish the convexity of the Lavery integral. It may be viewed as the extension of the  $L_1$  vector norm to a norm on the linear space  $F''$  of second derivatives of spline functions:

$$e(f) = \|f''\|_1.$$

The following generic seminorm properties are easily verified for the piecewise linear functions in  $F''$ ,

$$\begin{aligned} \|f''\|_1 = 0 &\iff f'' = 0 \\ |\alpha| \|f''\|_1 &= \|\alpha f''\|_1 \end{aligned}$$

along with the triangle inequality,

$$\|f^{(1)''} + f^{(2)''}\|_1 \leq \|f^{(1)''}\|_1 + \|f^{(2)''}\|_1.$$

Suppose the two spline functions  $f^{(1)}, f^{(2)}$  are actually two interpolating spline functions and, therefore, in the affine space  $S$ . Then their mean is again in  $S$  and, from the triangle inequality,

$$\left\| \frac{f^{(1)''} + f^{(2)''}}{2} \right\| \leq \frac{\|f^{(1)''}\| + \|f^{(2)''}\|}{2}$$

In terms of Lavery integrals,

$$\begin{aligned} &e\left(\frac{m_0^{(1)} + m_0^{(2)}}{2}, \dots, \frac{m_n^{(1)} + m_n^{(2)}}{2}\right) \\ &\leq \frac{e(m_1^{(1)}, \dots, m_n^{(1)}) + e(m_1^{(2)}, \dots, m_n^{(2)})}{2}, \end{aligned}$$

which establishes convexity. This leads immediately to

**Observation B:** *Any positive linear combination of Lavery splines for the same interpolation problem, — in particular, their mean —, is again a Lavery spline for this problem.*

#### 2.3.2 Uniqueness of Inflections

Contrary to the Holladay functional, which is strictly convex, the Lavery integral is not. As a result, uniqueness does not follow and, in fact, does not hold, as an example in Sec. 2.3.4 will show. Such minima of a convex function, however, must form a convex set.

Note that, in general,

$$\int \frac{|f^{(1)''}(x)| + |f^{(2)''}(x)|}{2} dx \geq \int \frac{|f^{(1)''}(x) + f^{(2)''}(x)|}{2} dx.$$

If both  $f^{(1)}$  and  $f^{(2)}$  are Lavery splines for the same interpolation problem, then so is their mean, and all three functions

$$f^{(1)''}, f^{(2)''}, \frac{f^{(1)''} + f^{(2)''}}{2}$$

return the same optimal value for their Lavery integrals. Thus equality holds in the above relation. This implies that both  $f^{(1)''}(x)$  and  $f^{(2)''}(x)$  have the same sign pattern:

$$f^{(1)''}(x) \geq 0 \text{ if and only if } f^{(2)''}(x) \geq 0.$$

This can be rephrased as

**Observation C:** *Two Lavery splines for the same optimization problem share essentially the same inflections: if one of them has an inflection point at  $x = \hat{x}$ , then so has the other unless it is linear at this point.*

### 2.3.3 Free Ends of Lavery Splines

Here we will examine the free ends of Lavery splines, in particular, the cubic polynomial  $f_1(x)$  and the corresponding first summand  $e(f_1)$  of the Lavery integral  $e(f)$ . As the end slope  $m_0$  may vary freely, it must optimize  $e(f_1)$  while keeping the slope  $m_1$  fixed. This fact determines the behavior of Lavery splines at free ends. As seen in the previous section, the first summand  $e(f_1)$  is a convex function in the variables  $m_0, m_1$ . If the slope  $m_1$  is held fixed,  $e(f_1)(m_0)$  is convex as a function in  $m_0$  alone. If the fixed slope equals the straight-line slope, then  $m_0 = m_1 = M_1$  obviously represents the optimal value for  $m_0$ , since  $f_1(x)$  in that case is a straight line with  $e(f_1) = 0$ . We suppose, therefore, that  $m_1 \neq M_1$ , and we examine the case, that  $f_1(x)$  has an inflection  $\hat{x}$  in the interval between the first two knots,  $x_0 \leq \hat{x} \leq x_1$ .

Consider the function

$$\tilde{e}(f_1)(m_0) = \frac{U_1^2 + V_1^2}{D_1},$$

where  $\hat{m}$  as well as  $U_1, V_1$  also depend on the variable  $m_0$ . Clearly, the absolute value of  $\tilde{e}(f_1)$  is given by  $e(f_1)$ . Note that

$$\begin{aligned} \frac{\partial \tilde{e}(f_1)}{\partial m_0} &= \frac{V_1^2 - U_1^2 - 6V_1U_1}{D_1^2} = \hat{\lambda}_1^2 - \hat{\mu}_1^2 - 6\hat{\lambda}_1\hat{\mu}_1 \\ &= 6\hat{\lambda}_1^2 - 4\hat{\lambda}_1 - 1, \end{aligned}$$

in view of (10), and

$$\frac{\partial U_1}{\partial m_0} = -2, \quad \frac{\partial V_1}{\partial m_0} = -1 \quad \frac{\partial D_1}{\partial m_0} = -3.$$

Solving the quadratic equation for  $\hat{\lambda}_1$ , and taking into account that  $0 \leq \hat{\lambda}_1 \leq 1$ , yields

$$\begin{aligned} (17) \quad \hat{\lambda}_1 &= \frac{2 + \sqrt{10}}{6} = 0.860378, \\ \hat{\mu}_1 &= \frac{4 - \sqrt{10}}{6} = 0.139622. \end{aligned}$$

The value of  $m_0$  for which this value for  $\hat{\lambda}_1$  is realized can be inferred from the general definition (15) of an inflection, giving the barycentric coordinate  $\hat{\lambda}_1$  in terms of the slopes  $m_0, m_1$ , as follows:

$$\hat{\lambda}_1(6M_1 - 3m_0 - 3m_1) = 3M_1 - m_0 - 2m_1,$$

which — for the particular value (17) of  $\hat{\lambda}_1$  — yields the corresponding end-slope

$$(18) \quad m_0 = \frac{10 - \sqrt{10}}{5} M_1 - \frac{5 - \sqrt{10}}{5} m_1.$$

This value for  $m_0$  represents a locally unique stationary value of  $\tilde{e}(f_1)(m_0)$ .

Now  $\tilde{e}(f_1)(m_0) = 0$  would imply  $e(f_1)(m_0) = 0$ , and consequently linearity, that is,  $m_0 = m_1 = M_1$ , which has been ruled out. By continuity,  $\tilde{e}(f_1)(m_0)$  is either always positive or always negative, — in other words, either

$$e(f_1)(m_0) = +\tilde{e}(f_1)(m_0)$$

or

$$e(f_1)(m_0) = -\tilde{e}(f_1)(m_0).$$

This implies that the value (18) for  $m_0$  is also a locally unique stationary value of  $e(f_1)(m_0)$ . In view of the convexity of this function, it is also its minimizer. At the last end we differentiate

$$(19) \quad \hat{e}(f_n) = \frac{U_n^2 + V_n^2}{3D_n}$$

and find

$$\frac{\partial \hat{e}(f_n)}{\partial m_n} = \frac{U_n^2 - V_n^2 - 6U_nV_n}{D_n^2} = 6\hat{\mu}_n^2 - 4\hat{\mu}_n - 1.$$

Symmetrically, we thus have

$$\hat{\mu}_n = \hat{\lambda}_1, \quad \hat{\lambda}_n = \hat{\mu}_1.$$

The equation

$$\hat{\mu}_n(6M_n - 3m_{n-1} - 3m_n) = 3M_n - 2m_{n-1} - m_n$$

thus yields a symmetric relationship to (18):

$$(20) \quad m_n = \frac{10 - \sqrt{10}}{5} M_n - \frac{5 - \sqrt{10}}{5} m_{n-1}.$$

This establishes

**Observation D:** *The free ends of Lavery splines are either linear functions or they contain an inflection in the interior of their interval of definition. The locations  $\hat{x}_1, \hat{x}_n$  of these inflections are universally given, respectively, by*

$$\begin{aligned}\hat{x}_0 &= \frac{2+\sqrt{10}}{6}x_0 + \frac{4-\sqrt{10}}{6}x_1, \\ \hat{x}_n &= \frac{4-\sqrt{10}}{6}x_{n-1} + \frac{2+\sqrt{10}}{6}x_n.\end{aligned}$$

Observation D enables us to determine universal values for partial Lavery integrals at the end-intervals of Lavery splines. By (15), (16), (18), and again by (20), which implies

$$\hat{\mu}_1^2 + \hat{\lambda}_1^2 = \hat{\lambda}_n^2 + \hat{\mu}_n^2 = \frac{10 - \sqrt{10}}{9},$$

we find

$$e(f_1) = \frac{10 - \sqrt{10}}{27} |D_1|, \quad e(f_n) = \frac{10 - \sqrt{10}}{27} |D_n|.$$

Substituting for  $m_0$  and  $m_n$ , respectively in  $D_1 = 6M_1 - 3m_0 - 3m_1$  and  $D_n = 6M_n - 3m_{n-1} - 3m_n$ ,

$$D_1 = \frac{3\sqrt{10}}{5}(M_1 - m_1), \quad D_n = \frac{3\sqrt{10}}{5}(M_n - m_{n-1}),$$

giving rise to

**Observation E:** *A necessary, but far from sufficient, condition for a spline function  $f$  to be a Lavery spline is that*

$$\begin{aligned}e(f_1) &= \frac{2}{9}(\sqrt{10} - 1) |M_1 - m_1|, \\ e(f_n) &= \frac{2}{9}(\sqrt{10} - 1) |M_n - m_{n-1}|.\end{aligned}$$

### 2.3.4 Examples of Non-unique Lavery Splines

In this section, we present an example in which the Lavery splines are not unique. Consider the three points (Figs. 10, 11, 12):

$$(21) \quad \begin{aligned}P_0 &= (x_0, y_0) = (-1, -1) \\ P_1 &= (x_1, y_1) = (0, 0) \\ P_2 &= (x_2, y_2) = (+1, -1).\end{aligned}$$

The associated interpolating cubic spline functions  $f(x)$  are then defined by their slopes

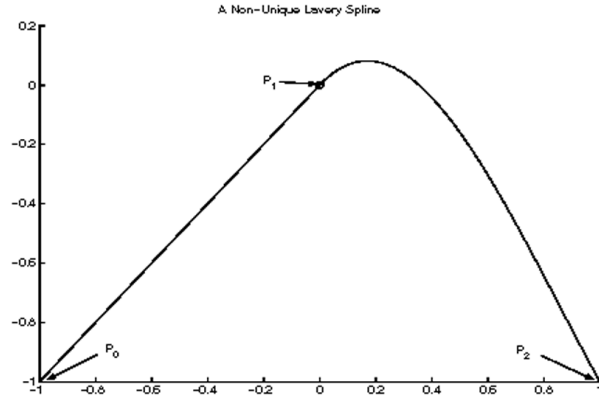
$$m_0, m_1, m_2$$

at these points. There are two subintervals with cubic polynomials

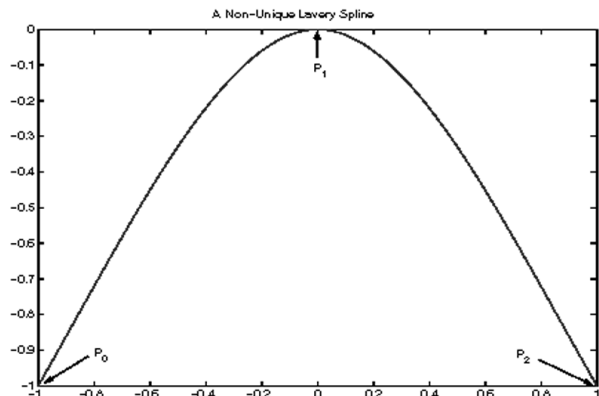
$$f_1(x), f_2(x),$$



**Fig. 10.** One of three Lavery splines for the same interpolation problem.



**Fig. 11.** Second of three Lavery splines for the same interpolation problem.



**Fig. 12.** Third of three Lavery splines for the same interpolation problem.

each of them a free end. This determines the coordinates  $\hat{x}_1, \hat{x}_2$  at inflections:

$$\hat{x}_1 = -\frac{2+\sqrt{10}}{6}, \hat{x}_n = +\frac{2+\sqrt{10}}{6}.$$

Note that both partial Lavery integrals are end-integrals. For an interpolating spline function  $f(x)$  to be a Lavery spline it will be necessary by Observation E that

$$e(f_1) = \frac{2}{9}(\sqrt{10}-1)|1-m_1|,$$

$$e(f_2) = \frac{\sqrt{10}-1}{9}|-1-m_1|$$

so that

$$e(f) = \frac{2}{9}(\sqrt{10}-1)(|1-m_1| + |-1-m_1|).$$

Clearly

$$(22) \quad -1 \leq m_1 \leq 1$$

implies

$$|1-m_1| + |-1-m_1| = |(1-m_1) + (-1-m_1)| = 2|m_1| \leq 2,$$

whereas either  $m_1 < -1$  or  $m_1 > 1$  would imply

$$|1-m_1| + |-1-m_1| = 2 \geq 2|m_1|.$$

Thus condition (22) characterizes all Lavery splines for the example.

Consider now any slope  $m_1$  from  $-1$  through  $+1$ . For  $m_1 = -1$ , we have by (18), (20), and in view of  $M_1 = 1$ ,  $M_2 = -1$ :

$$m_0 = \frac{+15-2\sqrt{10}}{5}, m_2 = -1.$$

Symmetrically, we find for  $m_1 = +1$  that

$$m_0 = +1, m_2 = \frac{-15+2\sqrt{10}}{5}.$$

These slopes, respectively, determine the two extreme Lavery splines, because each partial function is at a free end, and is optimized according to Observation D in the previous section. The two resulting Lavery splines are extreme in that each has a straight line segment as a partial function.

Using Hermite's formula (6) and substituting for  $x$ , we find for the choice  $m_1 = -1$ ,

$$f_1(x) = -\frac{2\sqrt{10}}{5}x^3 - \frac{10+2\sqrt{10}}{5}x^2 - x$$

$$f_2(x) = -x.$$

For  $m_1 = +1$ , the resulting Lavery spline is the symmetric image of the previous one:

$$f_1(x) = +x$$

$$f_2(x) = -\frac{2\sqrt{10}}{5}x^3 - \frac{10+2\sqrt{10}}{5}x^2 + x.$$

Both splines are shown in Figs. 10, 11. The self-symmetric spline from the choice  $m_1 = 0$  is shown in Fig. 12. In the latter case,

$$m_0 = \frac{10-\sqrt{10}}{5} = -m_2,$$

and

$$f_1(x) = -\frac{\sqrt{10}}{5}x^3 - \frac{5+\sqrt{10}}{5}x^2$$

$$f_2(x) = +\frac{\sqrt{10}}{5}x^3 - \frac{5+\sqrt{10}}{5}x^2.$$

This Lavery spline is the mean of the two splines with linear free ends. This is an instance of observation B about positive linear combinations of Lavery splines in Sec. 2.3.1.

## 2.4 Computing Univariate Lavery Splines

We now turn our attention to the computation of Lavery splines. The commonly used approach (see Lavery [11,12]) is to minimize the Lavery integral in discretized form, say,

$$(23) \quad \int_{x_0}^{x_n} |f''(x)| dx = \sum_i \frac{\Delta_i}{k_i} \sum_{k=1}^{k_i} |f''(x_{i-1} + \frac{k-1}{k_i} \Delta_i)|,$$

where the integrand is sampled in each subinterval  $[x_{i-1}, x_i]$  at  $k_i$  equidistant points. Now

$$x_{i-1} + \frac{k-1}{k_i} \Delta_i = \frac{(k_i - k + 1)x_{i-1} + (k-1)x_i}{k_i}$$

$$= \lambda_{i,k} x_{i-1} + \mu_{i,k} x_i, \lambda_{i,k} + \mu_{i,k} = 1.$$

Thus

$$\int_{x_0}^{x_n} f''(x) dx = \sum_i \frac{\Delta_i}{k_i} \sum_{k=1}^{k_i} |\lambda_{i,k} U_i + \mu_{i,k} V_i|$$

$$= \sum_i \frac{\Delta_i}{k_i} \sum_{k=1}^{k_i} |\lambda_{i,k} (3M_i - 2m_{i-1} - m_i) + \mu_{i,k} (3M_i - m_{i-1} - 2m_i)|$$

The Lavery integral is discretized as a sum of absolute values of the variables  $m_i$  of the minimization. Minimizing such an expression is a Linear Programming problem. Many satisfactory methods for

solving it are available, such as the Simplex Method or Interior Point Methods. The accuracy of the discretization increases with the number of sample points, but computational effort increases accordingly.

For that reason, and also to motivate an analogous approach in the bivariate case, we are proposing to minimize a different approximation to the Lavery integral, one that takes advantage of the ease of computation offered by energy minimization.

By the mean value theorem of integral calculus, there exist arguments  $u_i$  such that

$$\int_{x_{i-1}}^{x_i} f_i''(x)^2 dx = \Delta_i (f_i''(u_i))^2, \quad i=1, \dots, n.$$

we then propose to approximate the Lavery integral by the following Riemann sum:

$$(24) \quad \int_{x_0}^{x_n} |f''(x)| dx \approx \sum_i \Delta_i |f_i''(u_i)|.$$

In contrast to the approximation (23) by discretization, this approximation does not offer the option of further refinement, unless the interpolation problem itself is changed by adding additional knots and ordinates. In a sense, it approximates the Lavery paradigm itself rather than the Lavery integral. We still use, however, the term “approximate Lavery integral” for our proposed approximation (24).

#### 2.4.1 An Iterative Algorithm for Approximate Lavery Splines

For the purpose of computation, we rewrite the non-zero terms in (46)

$$\begin{aligned} \Delta_i \sqrt{f_i''(u_i)^2} &= \sqrt{\Delta_i} \frac{\Delta_i f_i''(u_i)^2}{\sqrt{\Delta_i f_i''(u_i)^2}} \\ &= \frac{\sqrt{\Delta_i}}{\sqrt{\int_{x_{i-1}}^{x_i} f_i''(x)^2 dx}} \int_{x_{i-1}}^{x_i} f_i''(x)^2 dx. \end{aligned}$$

This expression suggests an iterative approach. Starting with the classic spline  $f^{(0)}(x)$ , a sequence of interpolating spline functions is generated in hopes to converge towards the approximate Lavery integral (24)

$$f^{(0)}(x), f^{(1)}(x), \dots, f^{(l)}(x), \dots$$

with associated partial Holladay integrals

$$E_{f_i^{(l)}} = \int_{x_{i-1}}^{x_i} \left( f_i^{(l)''}(x) \right)^2 dx, \quad i=1, \dots, n.$$

At each step  $l = 0, 1, 2, \dots$ , weights

$$w_i^{(l)} = \sqrt{\frac{\Delta_i}{E(f_i^{(l)})}}, \quad i=1, \dots, n,$$

are introduced. Given the function  $f^{(l)}$ , the subsequent function  $f^{(l+1)}$  is determined as the solution to the following minimization problem:

$$(25) \quad \min \sum_i w_i^{(l)} \int_{x_{i-1}}^{x_i} \left( f_i^{(l+1)''}(x) \right)^2 dx.$$

This approach raises the question what to do if  $E(f_i^{(l)})$  vanishes as the corresponding weight is then not defined?

Simply ignoring such terms may prematurely lock in straight line segments between knots. What first comes to mind is to specify a limit  $\varepsilon > 0$  and boost lower values of  $E(f_i^{(l)})$  to this level. A more diligent procedure might be to start with all weights at value 1, — the weight setting that yields the initial classical spline according to Holliday’s observation —, and then progressively increase the use of weights. Such strategies remain to be explored.

In general, adding up partial Holliday integrals, each with weight, say,  $w_i$  leads again to an expression of the energy of a physical structure: a collection of thin beams of different thicknesses given by  $w_i$ , respectively, and welded together at knots. Minimizing this energy expression requires an adjustment to the linear system (12).

The weighted energy of the partial spline functions  $f_i(x)$  is just the product of the straight Holliday integral and the respective weight:

$$E_w(f_i) = w_i E(f_i).$$

The total weighted energy is thus given by

$$E_w(f) = \sum_i E_w(f_i) = \sum_i w_i E(f_i).$$

This means that in the expression (11) for  $E(f)$ , the factor

$$\frac{4}{\Delta_i}$$

is affected as it is multiplied by  $w_i$ . In other words, 1the substitution,

$$\frac{1}{\Delta_i} \rightarrow \frac{w_i}{\Delta_i},$$

transforms the linear system (12) for minimizing  $E(f)$  into the linear system for minimizing  $E_w(f)$ :

$$\begin{aligned}
(26) \quad & \frac{2w_1}{\Delta_1} m_0 + \frac{w_1}{\Delta_1} m_1 = \frac{3w_1}{\Delta_1} M_1 \\
& \frac{w_1}{\Delta_1} m_0 + \left( \frac{2w_1}{\Delta_1} + \frac{2w_2}{\Delta_2} \right) m_1 + \frac{w_2}{\Delta_2} m_2 = \frac{3w_1}{\Delta_1} M_1 + \frac{3w_2}{\Delta_2} M_2 \\
& \frac{w_2}{\Delta_2} m_1 + \left( \frac{2w_2}{\Delta_2} + \frac{2w_3}{\Delta_3} \right) m_2 + \frac{w_3}{\Delta_3} m_3 = \frac{3w_2}{\Delta_2} M_2 + \frac{3w_3}{\Delta_3} M_3 \\
& \quad \quad \quad \dots = \dots \\
& \frac{w_n}{\Delta_n} m_{n-1} + \frac{2w_n}{\Delta_n} m_n = \frac{3w_n}{\Delta_n} M_n
\end{aligned}$$

Note that the weighted classic splines are, in general, not twice differentiable at the knots. However, due to the fact that the first and the last equations have common factors  $w_0$ ,  $w_n$ , respectively, they are equivalent to the corresponding equations in the Holladay system (12): free ends thus have zero second derivatives in the weighted case, too. Again, this is to be expected from Physics.

Note also that some commonly used methods for determining classic splines such as B-splines or linear systems formulated in terms of second derivatives at knots do not carry over to the weighted case. However, the above linear system is still “banded”, and many excellent methods are known for its solution. To solve this system we used the venerable Gauss-Seidel method, not just for ease of programming, but also because it seems to work for bivariate weighted splines. Its advantage lies in the fact that the matrix of the linear system need not be changed and can be read, so to speak, in sequence. This is important for the very large systems likely to arise in the bivariate case. In the univariate case, the convergence behavior is well understood (See Varga [22]). Using an iterative method for solving the class of linear systems above will result in a two-tiered iteration procedure: an “outer” iteration, developing new sets of weights, and an “inner” iteration, solving the resulting linear system. Such procedures can be “balanced”, that is, the inner iteration may be terminated at a lower accuracy level during the early stages of the outer iteration and may be carried to a higher level of accuracy as the outer method approaches convergence. This is an added advantage of an iterative method for solving the linear systems at hand.

Note finally, that the approximate Lavery method proposed here will definitely not converge to the optimal Lavery integral. This is because the approximate solutions are based on weighted classic splines, and therefore have vanishing second derivatives at the free

ends. The second derivatives of Lavery splines, on the other hand, assume their inflections in interiors of the end interval (See Observation D).

**Observation F:** *Unless a free end function of an approximate Lavery spline is linear, it disagrees with the corresponding end function of the true Lavery spline in that the latter has an inflection in the interior of at least one end interval, whereas the former assumes its inflection at the end knot.*

The justification for introducing the approximate Lavery splines concept lies in its computational ease and the fact that it appears to retain the anti-oscillatory dynamic of the original Lavery concept. In fact, the examples in Figs. 5, 6, and 8 were calculated using the approximate algorithm outlined in this section.

### 3. Bivariate Spline Interpolation

The development of bivariate splines parallels that of univariate ones to a large extent. For both, the first step is to specify a linear space  $F$  of bivariate spline functions  $f(x, y)$ , defined on a polygonally partitioned region

$$\Omega$$

In view of our emphasis on the TIN framework of triangulations, however, we choose for our discussion a rectangular region  $\Omega$  and an irregular set of “knots” in  $\Omega$ ,

$$v_i = (x_i, y_i) \in \Omega, i = 1, \dots, n.$$

An interpolation problem results, if “elevations”  $z_i$  are specified at the vertices  $v_i = (x_i, y_i)$ .

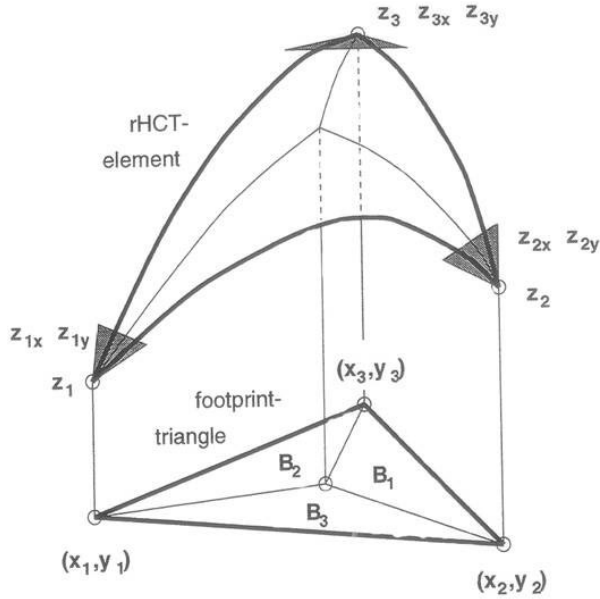
The region  $\Omega$  is then partitioned into triangles

$$t_k.$$

The vertices

$$(x_{i_{k,1}}, y_{i_{k,1}}), (x_{i_{k,2}}, y_{i_{k,2}}), (x_{i_{k,3}}, y_{i_{k,3}})$$

of any such triangle  $t_k$  are among the specified knots, and those are the only knots in this triangle. In this fashion, a “triangulation” of the region  $\Omega$  is obtained. Spline functions are then defined by installing partial functions



**Fig. 13.** Reduced Hsieh-Clough-Tocher (rHCT) element.  $z_{ix}$ ,  $z_{iy}$  denote the partial slopes  $zx_i$ ,  $zy_i$  (27).

$$f_k(x, y), \quad k = 1, \dots, m$$

in their respective triangles  $t_k$  of the triangulation.

Following C. Lawson [14,15], we install above each triangle  $t_k$  a reduced Hsieh-Clough-Tocher (rHCT) element. The reduced Hsieh-Clough-Tocher (rHCT) element, displayed in Fig. 14, provides a means for representing smooth TIN surfaces (see Lawson [14,15]). The rHCT element is defined as a function over a triangle in the  $x, y$ -plane, and is fully determined by the elevations and the partial slopes or derivatives at triangle corners. The triangle is divided into three subtriangles with their common vertex at the centroid of the three corners of the full triangle. In each of the three subtriangles the rHCT function is a bivariate cubic function. Together, these functions form a  $C^1$  function on the full triangle (Fig. 13). This smooth rHCT function is furthermore constrained by the linear perpendicularity condition on the derivatives orthogonal to the outside boundary edges of the triangular element: along each such edge, these derivatives are required to be linear functions. It follows that the orthogonal derivatives along an edge are fully determined by their values at the ends of the edges. These values are inherent to both triangles adjacent to the edge so that the orthogonal derivatives coincide when calculated within each of the two triangles independently. The above condition thus provides the key property ensuring smoothness of a rHCT surface.

As will be seen in Sec. 3.1.2, the restriction inherent in the linear perpendicularity condition may be more severe than commonly expected.

To this end, “elevations”

$$z_i$$

and “partial slopes”

$$(27) \quad zx_i, zy_i, \quad i = 1, \dots, n,$$

are prescribed at the knots  $v_i$ . As the vertices of each triangle  $t_k$  are among the specified knots, the corresponding elevations

$$z_{i_{k,1}}, z_{i_{k,2}}, z_{i_{k,3}},$$

and partial slopes

$$zx_{i_{k,1}}, zy_{i_{k,1}}, zx_{i_{k,2}}, zy_{i_{k,2}}, zx_{i_{k,3}}, zy_{i_{k,3}}.$$

furnish the parameters necessary for defining the rHCT element  $f_k(x, y)$ . As pointed out above in the current section, each specification of elevations and partial slopes at the knots yields a  $C^1$  function  $f(x, y)$ , which also represents a smooth surface over the region of definition  $\Omega$ . All such functions, — relating to the given triangulation of  $\Omega$  —, constitute a linear space of “spline functions” and their respective derivative linear spaces

$$F, F_x, F_y, F_{xx}, F_{xy}, F_{yy}.$$

If a specific set of elevations  $z_i$  is to be interpolated, then those spline functions which meet these elevations constitute an affine space of “interpolating spline functions” and its derivative affine spaces

$$S, S_x, S_y, S_{xx}, S_{xy}, S_{yy}.$$

Note, again, that second derivatives are not fully defined for rHCT elements. Indeed, along edges separating two subtriangles of an rHCT element as well as along edges separating two triangles of the triangulation there may be two different values in contention depending from which side a point on such an edge is approached. For purposes of integration, this is not an issue since the discrepancies are restricted to a set of measure zero, so that the integrals considered below are still well defined.



### 3.1 Paradigms for Non-Oscillatory Bivariate Splines

For the purpose of interpolation, the elevations  $z_i$  at knots are specified, and the task at hand is to determine the partial slopes  $z_{x_i}, z_{y_i}$  so as to arrive at a “best” interpolating spline function. Such an optimal spline function is then referred to as a “(bivariate) spline”. Obviously, there are various kinds of splines depending on the choice of the linear space of spline functions and the choice of the optimization criterion.

Paralleling the classical univariate approach, “thin plate” minimization exemplifies the “classical” approach, and the resulting spline will be referred to as the “classical spline”, the functional requiring minimization being

$$(28) \quad \tilde{E}(f) = \int_{\Omega} (f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2) dx dy.$$

An approach to computing the surface energy formula has been announced by McClain and Witzgall [16] for rHCT element. A newer version of the report is currently being prepared by McClain, Witzgall, and Gilsinn [17].

The authors believe that, contrary to univariate classical splines, bivariate classical splines as defined above are, in general, not in class  $C^2$ , that is, twice differentiable. Nevertheless they are “stiff” and thus subject to the dreaded spurious oscillations and Gibbs phenomena. This can be seen along edges of buildings in the urban data set of Baltimore, MD, given in Fig. 14. Lavery and Gilsinn [13] were able to address this problem by extending his paradigm for univariate nonoscil-

latory splines to the bivariate case by minimizing the functional

$$(29) \quad e(f) = \int_{\Omega} (|f_{xx}| + 2|f_{xy}| + |f_{yy}|) dx dy.$$

Again, the squares of second derivatives in an energy integral, — here  $\tilde{E}(f)$  as defined in (28) —, are replaced by absolute values. We will refer to the so defined splines also a “(bivariate) Lavery splines”.

In view of the non-uniqueness result for univariate Lavery splines, the bivariate Lavery splines are expected not to be unique either, as will be discussed below. Therefore, the Lavery paradigm typically includes a regulatory term such as adding a small multiple of  $f'(x, y)^2$ . The resulting bivariate Lavery splines have produced oscillation-free interpolations and approximations in a large array of applications. They are computed by discretizations leading to very large linear programming problems.

For large surface generation tasks based on interpolation, the resulting computational effort may be prohibitive. Moreover, the bivariate Lavery integral (29) functional is not invariant under rotation of the coordinate system. The authors have therefore proposed an alternate extension of the univariate Lavery paradigm to the bivariate case:

$$(30) \quad \tilde{e}(f) = \int_{\Omega} \sqrt{f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2} dx dy.$$

Note, that this paradigm is indeed an extension of the univariate Lavery paradigm. Indeed,

$$\int |f''(x)| dx = \int \sqrt{f''(x)^2} dx,$$

so that the integrand of the univariate Lavery integral may also be interpreted as the squareroot of an energy, — in this case, the elastic energy of a thin beam. Replacing that energy expression by that for a thin plate, yields the alternate extension.

#### 3.1.1 Properties of the Alternative Bivariate Paradigm

The differential expression

$$(f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2) dx dy$$

is readily seen to be rotation invariant, as is to be expected, given its physical meaning as the integrand for the representation (28) of the thin plate energy. It follows that every differential expression of the form



Fig. 14. Urban scene (Baltimore): Gibbs phenomena are visible.

$$\Phi(f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2)dxdy$$

is also rotation invariant. This leads to the

**Observation G:** *The Alternate Bivariate Functional  $\bar{e}(f)$  is invariant under a rotation in the  $x, y$ -plane.*

Let  $f = f(x, y)$ ,  $g = g(x, y) \in S$  be two interpolating spline functions. Invoking the general vector inequality  $\|u\| + \|v\| \geq \|u + v\|$ , we find

$$\begin{aligned} & \frac{\sqrt{f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2} + \sqrt{g_{xx}^2 + 2g_{xy}^2 + g_{yy}^2}}{2} \\ & \geq \sqrt{\left(\frac{f_{xx} + g_{xx}}{2}\right)^2 + 2\left(\frac{f_{xy} + g_{xy}}{2}\right)^2 + \left(\frac{f_{yy} + g_{yy}}{2}\right)^2}, \end{aligned}$$

so that

$$\begin{aligned} & \frac{1}{2} \int_{\Omega} \sqrt{f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2} dxdy + \frac{1}{2} \int_{\Omega} \sqrt{g_{xx}^2 + 2g_{xy}^2 + g_{yy}^2} dxdy \\ & \geq \int_{\Omega} \sqrt{\left(\frac{f_{xx} + g_{xx}}{2}\right)^2 + 2\left(\frac{f_{xy} + g_{xy}}{2}\right)^2 + \left(\frac{f_{yy} + g_{yy}}{2}\right)^2} dxdy, \end{aligned}$$

and in terms of the spline functionals,

$$\frac{\bar{e}(f) + \bar{e}(g)}{2} \geq \bar{e}\left(\frac{f + g}{2}\right).$$

Since  $S$  is an affine space, the mean  $(f + g)/2$  is again in  $S$ . Based on the above inequality, we thus have

**Observation H:** *The alternate Bivariate functional (30) is a convex functional. In particular, as expressed in terms of the partial slopes,*

$$\bar{e}(f) = \bar{e}(zx_1, zy_1, \dots, zx_n, zy_n),$$

*it is a convex function in terms of these partial slopes.*

### 3.1.2 A Related Non-Uniqueness Result

We do not know at this point, whether the convexity is moreover strict. We expect that it is not. Strict convexity would imply uniqueness of minima and, again, we do not believe that this is the case. However, an example for such non-uniqueness is eluding us for our specific space of rHTC spline functions.

For different bivariate functions, however, it is possible to establish non-uniqueness of a minimum. Consider a function

$$s(x, y), -1 \leq x \leq +1, 0 \leq y \leq 1$$

representing a rectangular strip bent only in the direction of the  $x$ -axis. In the direction of the  $y$ -axis, the function  $s(x, y)$  is constant (Fig. 15), in particular, for all  $y$ ,

$$s(-1, y) = -1, s(0, y) = 0, s(+1, y) = +1.$$

The function  $s(x, y)$  thus interpolates the values  $-1, 0, +1$  along the contour lines in the direction of the  $y$ -axis at the arguments  $x = -1, 0, +1$ . Given any Lavery spline  $f(x)$  considered in Sec. 2.3.3 as interpolating the example (21). Then it is readily seen, that any function

$$s(x, y) = f(x)$$

minimizes both bivariate functionals:  $e(s)$ ,  $\bar{e}(s)$ . For functions corresponding to strips bent in only the strip direction, these bivariate functionals have thus multiple minima. This makes a strong case for non-uniqueness bivariate splines minimizing the functionals (29) and (30). However, we were not able to pinpoint non-uniqueness for the rHTC based spline functions considered in this section. In particular, we were surprised to realize that the function  $s(x, y)$  cannot be generated using rHTC elements. As this fact points to an important limitation of using rHTC elements, we formulate the

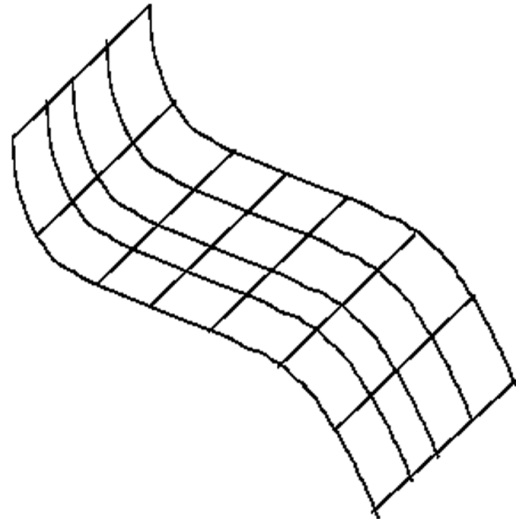


Fig. 15. Non-uniqueness of a minimum for a bivariate function.

**Observation J:** *The surface function of a rectangular strip bent only in the strip direction cannot be represented by means of rHTC elements.*

### 3.2 Computing Bivariate Splines Based on an Alternate Paradigm

The proposed alternate functional (53) is still expensive to minimize. In analogy to the procedure described in Sec. 2.4.1, we propose to approximate that functional as follows. Let

$$\Delta_k$$

denote the area of the triangle  $t_k$  of the triangulation. Then the integral of each partial spline function or element  $f_k = f_k(x, y)$  is individually approximated.

$$\begin{aligned} & \int_{t_k} \sqrt{f_{kxx}^2 + 2f_{kxy}^2 + f_{kyy}^2} \, dx dy \\ &= \frac{\sqrt{\Delta_k}}{\sqrt{\int_{t_k} (f_{kxx}^2 + 2f_{kxy}^2 + f_{kyy}^2) \, dx dy}} \int_{t_k} (f_{kxx}^2 + 2f_{kxy}^2 + f_{kyy}^2) \, dx dy. \end{aligned}$$

This approximation suggests an iterative procedure based on minimizing weighted energy expressions: Starting with the classic spline  $f^{(0)}(x, y)$ , a sequence of interpolating spline functions is generated in hopes to converge towards the alternate Lavery integral (30),

$$f^{(0)}(x, y), f^{(1)}(x, y), \dots, f^{(l)}(x, y), \dots$$

with associated partial integrals

$$\tilde{E}(f_k^{(l)}) = \int_{t_k} (f_{kxx}^{(l)2} + 2f_{kxy}^{(l)2} + f_{kyy}^{(l)2}) \, dx dy, \quad k=1, \dots, m$$

At each step  $l = 0, 1, 2, \dots$ , weights

$$w_k^{(l)} = \sqrt{\frac{\Delta_k}{\tilde{E}(f_k^{(l)})}}, \quad k=1, \dots, m$$

are introduced. Given the function  $f^{(l-1)}$ , the subsequent function  $f^{(l)}$  is determined as the solution to the following minimization problem (compare (25)):

$$(31) \quad \min \sum_i w_i^{(l-1)} \int_{t_i} (f_{ixx}^{(l)2} + 2f_{ixy}^{(l)2} + f_{iyy}^{(l)2}) \, dx dy.$$

For the exploratory implementation reported here, the above minimization of a quadratic function in the partial slopes

$$zx_{i_{k1}}, zy_{i_{k1}}, zx_{i_{k2}}, zy_{i_{k2}}, zx_{i_{k3}}, zy_{i_{k3}}.$$

is again formulated as solving a linear system of equations in these variables, and Gauss-Seidel iteration provided a workable option. Just as discussed in the univariate case, there are two iterative processes, one “on top of” the other. The “inner iteration” aims to solve the linear system of equations in order to solve the intermediate minimization problem (31), while the “outer” iteration sequentially creates spline functions expected to converge toward a limit that approximates the alternate Lavery spline defined by the minimization of the alternate Lavery functional (30).

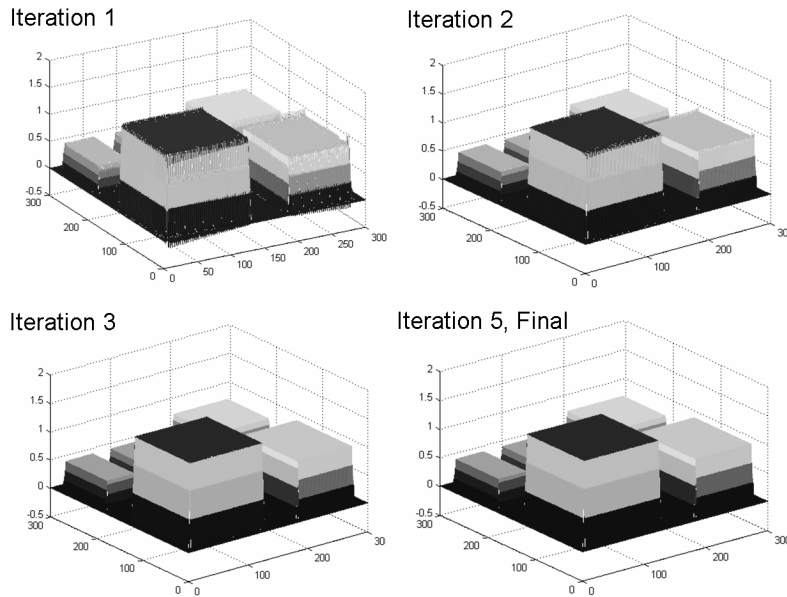
This approach raises the same issues as its parallel in the univariate case. The above iterative approximation based on consecutive minimization of weighted sums of integrals (31) cannot be expected to converge exactly to the alternate bivariate Lavery spline. The argument relies on Observation F pertaining to the univariate case. Each iterate  $f^{(l)}(x, y)$  minimizes the energy of a physical structure consisting of thin plates of various thicknesses, namely the weights. Any minimizing shape of that structure, however, is known to have some vanishing second derivatives at its boundary. Univariate Lavery splines do not have this property, and we do not expect bivariate Lavery splines to have this property, either.

Provisions need to be made should one of the integrals  $\tilde{E}(f_k^{(l-1)})$  vanish. This is a serious problem because very large or “infinite” weights will cause certain rHTC elements to remain stuck in linear form. In addition, limits of bivariate Lavery splines, in the original as well as the alternate formulation, may not be unique (see Sec. 3.1). These issues call for regularization procedures which are adequate for particular kinds of applications.

Indeed, spurious spikes were observed for the applications demonstrated here. The problem was solved by interspersing an averaging step, where the partial slopes were replaced by averages of neighboring slopes.

While our experiences with approximating alternate bivariate Lavery splines (Fig. 16) were encouraging, in particular, for very large data sets, much work needs to be done.

The accuracy performance of the Gauss-Seidel iteration is a major issue. Simple examples show that initial steps are not even contracting. When does contraction start? Does it keep contracting after that? Much research has been aimed at these questions. These analyses of the convergence properties of the Gauss-



**Fig. 16.** Example of 2-D Lavery splines.

Seidel iteration, however, provide estimates of the number of iterations required to solve linear equations to a specified accuracy which are much too large to be practical. To be practical, the number of “passes” through the entire system of linear equations on the order of 10 or 20, has to be sufficient.

In our experience, relatively few steps of the “outer” iteration were sufficient to provide qualitatively satisfactory results. However, we do not have formal proof of convergence.

Finally, and perhaps most importantly, a better understanding of the necessary regularization devices needs to be developed.

## 4. Conclusions

In this paper we have investigated some properties of non-oscillatory splines introduced by John Lavery [11,12]. These splines, called in this paper Lavery splines, minimize what we have termed the Lavery integral (4). We have seen that the minimizing spline for (4) does indeed model sharp edges and jumps in data without introducing the “Gibbs phenomenon” at the corners. We have shown that the Lavery integral and the associated Lavery splines satisfy a number of properties. First, we have shown that the Lavery integral of a cubic spline function is the absolute value of an alternating sum of inflection slopes and of the end

slopes. Next, we showed that any positive linear combination of Lavery splines for the same interpolation problem is again a Lavery spline for the same problem. Furthermore, two Lavery splines for the same optimization problem share essentially the same inflection points and that the free ends of Lavery splines are either linear functions or they contain an inflection point in the interval of definition.

Two algorithms for estimating Lavery splines have also been considered. The first algorithm introduced by John Lavery [11,12] reduces to solving a least absolute value minimization problem for which he used an interior point method for linear programming to obtain the minimum spline coefficients. The absolute value minimization was based on a “discretization” which leads to a very large number of variables as sufficiently smaller discretizations were considered. The extension of this method to bivariate Lavery splines also led to computationally intensive compute times even for moderate data sizes. The authors have introduced a modified approach based on an iterated weighted least squares algorithm. Although the minimizing spline this algorithm produces is not a Lavery spline it is an approximation that also produces sharp edges without the “Gibbs phenomenon.”

In the bivariate case the Lavery integral minimizes the integral of the sum of the absolute values of the second partial derivatives of the spline. Analogous to the univariate case we introduce an alternative variational

principle based on the integral of the square root of the sum of squares of the second partial derivatives of the bivariate spline. Whereas the Lavery integral in the bivariate case is not rotationally invariant, the alternative principle is. The alternative bivariate functional is shown to be a convex functional and, when expressed in terms of the partial slopes at the vertices of an underlying element, it is a convex function of these slopes. We have also shown that a rectangular strip bent only in the strip direction cannot be represented by means of a triangular rHTC element.

Analogous to the alternative weighted least squares algorithm introduced to compute the approximate univariate splines we introduce an alternative weighted least squares algorithm in the bivariate case. Although we have not introduced a convergence analysis of the alternative algorithms in this paper, computational experience has shown that the iterative, Gauss-Seidel based, algorithm used provides qualitatively satisfactory results in only a few iterations of the main loop.

## 5. References

- [1] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, *The Theory of Splines and Their Applications*, Academic Press, New York (1967).
- [2] P. J. Besl and N. D. McKay, A method for registration of 3-D shapes, *IEEE Trans. on Pattern Analysis and Machine Intelligence* **14**, 239-256 (1992).
- [3] A. K. Cline, Scalar-and Planar-Valued Curve Fitting Using Splines Under Tension, *Communications of the ACM* **17**, 218-220 (1974).
- [4] A. K. Cline, Algorithm 476 - Six subprograms for curve fitting using splines under tension, *Communications of the ACM* **17**, 220-223 (1974).
- [5] C. Delaunay, Sur la sphere vide, *Bull. Acad. Sci. USSR (VII)*, Classe Sci. Mat. 793-800 (1934).
- [6] D. E. Gilsinn, Constructing Sibson Elements for a Rectangular Mesh, NISTIR 6718, National Institute of Standards and Technology, Gaithersburg, MD (2001).
- [7] D. E. Gilsinn, M. McClain, and C. Witzgall, Using Nonoscillatory Splines to Model Urban Environments, *Geometric Modeling and Computing*: Seattle, 2003, M. L. Lucian and M. Neamtu, eds., Nashboro Press, Brentwood (2004) pp. 229-248.
- [8] D. E. Gilsinn and J. Lavery, Shape-preserving, Multiscale Fitting of Bivariate Data by Cubic L1 Smoothing Splines, *Approximation Theory X: Wavelets, Splines, and Applications*, C. K. Chui, L. L. Schumaker, and J. Stockler, eds., Vanderbilt University Press, Nashville, TN (2002).
- [9] D. E. Gilsinn and J. Lavery, Representation of Natural Terrain by Cubic L1 Splines, *Trends in Approximation Theory*, K. Kopotun, T. Lyche, and M. Neamtu eds., Vanderbilt University Press, Nashville, TN (2001).
- [10] D. E. Gilsinn and J. E., Lavery, "Representation of Urban Terrain by L1 Smoothing Splines", *Proceeding Tenth International Conference on Approximation Theory*, St. Louis, MO, March 26-29 (2001).
- [11] J. E. Lavery, Univariate cubic  $L_p$  splines and shape preserving, multiscale interpolation by univariate cubic  $L_1$  splines, *Comput. Aided Design* **17**, 319-336 (2000).
- [12] J. E. Lavery, Shape-preserving, multiscale fitting of univariate data by cubic  $L_1$  smoothing splines, *Comput. Aided Geom. Design* **17**, 715-727 (2000).
- [13] J. E. Lavery and D. E. Gilsinn, Multiresolution Representation of Urban Terrain By L1 Splines, L2 Splines and Piecewise Planar Surfaces *Proceeding 22nd Army Science Conference in Baltimore, MD*, 11-13 December (2000).
- [14] C. L. Lawson,  $C^1$ -compatible interpolation over a triangle, Technical Memorandum No. 33-770, Jet Propulsion Laboratory, Pasadena, California (1976).
- [15] C. L. Lawson, Software for  $C^1$  surface interpolation, *Mathematical Software III*, *Proceedings of a Symposium conducted by the Mathematical Research Center, University of Wisconsin, Madison*, J. R. Rice, ed., Academic Press, New York (1977) pp. 161-194.
- [16] M. McClain and C. Witzgall, Surface Energy Formula for a Hsieh-Clough-Tocher Element, NISTIR 6246, National Institute of Standards and Technology, Gaithersburg, MD (1999).
- [17] M. McClain, C. Witzgall, and D. Gilsinn, A New Surface Energy Formula for a Hsieh-Clough-Tocher Element, NISTIR (in preparation), National Institute of Standards and Technology, Gaithersburg, MD.
- [18] B. Mandel, J. Bernal, and C. Witzgall, Contour-to-Grid Interpolation with Nonlinear Finite Elements: A Feasibility Study, ETL-0472, U. S. Corps of Engineers, Engineer Topographic Laboratories, Fort Belvoir, VA 22060-5546.
- [19] T. K. Peucker, R. J. Fowler, J. L. Little, and D. M. Mark, The Triangulated Irregular Network, *Proceedings of the Digital Terrain Model Symposium*, St. Louis, MO, 516-532 (1978).
- [20] D. G. Schweikert, An interpolation curve using a spline in tension, *Journal of Mathematics and Physics* **45**, 312-317 (1966).
- [21] J. Stoer and B. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, New York (1993).
- [22] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1962).

**About the authors:** *Christoph Witzgall holds the designation of Scientist Emeritus from the NIST Information Technology Laboratory (ITL). Although he retired from government service in 2003, he continues to serve as a guest researcher in the ITL Mathematical and Computational Sciences Division. Christoph has been associated with NIST since 1962, serving as Acting Chief of its Operations Research Division from 1979 to 1982. He received the Department of Commerce Silver Medal for meritorious Federal service. David E. Gilsinn is a mathematician with the Mathematical and Computational Sciences Division of ITL. Marjorie A. McClain is a mathematician with the Mathematical and Computational Sciences Division of ITL. The National Institute of Standards and Technology is an agency of the Technology Administration, U.S. Department of Commerce.*